

Towards a Behaviourally Based Performance Appraisal System for Academic Staff

J.S. POUNDER
Lingnan College Hong Kong

This paper describes an academic performance appraisal initiative undertaken by the author in conjunction with academic staff and students of the Department of Management of the Hong Kong Polytechnic. The initiative involved the development of a set of behaviourally based performance rating scales and the design of a multi-appraisal method. In addition, a strategy for implementation was proposed. It is argued that the current emphasis on performance measurement in Hong Kong higher education has given particular relevance to this initiative.

Introduction

There is widespread interest in performance measurement in higher education. Quality is a common term in academic circles these days. Consistent with this trend, in Hong Kong, the University and Polytechnic Grants Committee (UPGC) recently completed a research assessment of all UPGC funded institutions, and at the time of writing, is on the point of embarking on a similar assessment of teaching. Against that background, this paper describes an academic staff performance appraisal initiative undertaken in 1990 by its author, in collaboration with staff and students of the Hong Kong Polytechnic Department of Management.

A performance appraisal system should have both summative and formative objectives (Centra, 1982). As regards the former, a well designed appraisal instrument should provide an information source to those charged with making judgements relating to such issues as promotion, tenure and dismissal and reduce the subjective element in such decisions.

Equally, with regard to the latter, it is important for appraisal to be linked to development,

and in this respect, there is evidence to suggest that knowledge of appraisal results is likely to lead to improved performance on the part of those persons being appraised (Landy and Farr, 1983).

An important aim of the initiative was to devise a performance appraisal system that would lend itself to both objectives.

Specific Objective of the Performance Appraisal Initiative

In addition to the above, the selection of an instrument comprising Behaviourally Anchored Rating Scales (BARS) for development in the Department of Management of the Hong Kong Polytechnic was perceived as meeting the specific objective of laying a foundation for staff acceptance of the principle of academic performance appraisal.

That perception was based on management of change literature which consistently stresses the importance of participation and involvement as a means of reducing resistance to change (see, for example, Kotter and Schlesinger, 1979). It was felt that the participative nature of the BARS developmental procedure would be useful in a situation,

such as that prevailing in the Department of Management of the Hong Kong Polytechnic, where academic staff had little experience of systematic appraisal.

However, for appraisal to be effective, general agreement on the constituents of good performance is necessary. A literature review suggested that in the academic world this is far from the case.

The Problem of Academic Priorities

There is an interesting quotation from Caplow and McGee (1961) contained in Greenaway and Harding (1978), which sheds light on the problem:

Although in most occupations men are judged by how well they perform their normal duties, the academic man is judged almost exclusively by his performance in a kind of part time voluntary job which he creates for himself. It is only a slight exaggeration to say that academic success is likely to come to the man who has learned to neglect his assigned duties in order to have more time and energy to pursue his private professional interests, (p. 40)

Whilst the above quotation represents an extreme view, it does raise the issue as to where the priorities of an academic should lie.

A 1976 survey conducted by Greenaway and Harding revealed a reluctance on the part of academic institutions in the U.K. to commit themselves to any declaration of academic competence. When asked to state their criteria of promotion through the salary bar, 43% of universities and 30% of the then polytechnics responded nothing more detailed than "competence in teaching, research, and administration".

Approximately 40% of university and 60% of polytechnic respondents offered no criteria and only 20% of universities and 10% of polytechnics offered a reasonably detailed statement. Greenaway and Harding's analysis indicated four main areas of competence used as criteria for promotion, namely, teaching, research, administration, and a general area including factors such as extent of cooperation and responsibility, level of initiative, and degree of commitment to the institution.

The findings of Greenaway and Harding indicated that, when pressed, institutions imply that there is an acceptable balance between teaching, research, and administration etc. However literature suggests a measure of disagreement as to the

importance of the various elements of the academic's job. Blackburn et al (1980) have noted that U.S. faculty members believe tenure and/or promotion to be dependent on scholarly productivity rather than teaching competence. Equally, they have stated that many academics do not see the administration role as being a significant factor in promotion. Additionally, they have noted the 'schizophrenic' attitude frequently displayed by institutions to the consulting activity. It is approved of when it brings fame to the institution yet disapproved of because it takes time away from the job the institution is paying the academic to perform.

A review of some pertinent statements made in the literature indicates the differing views on academic priorities.

Miller (1974) emphasises the teaching input as follows:

Classroom teaching for a large majority of colleges and universities is the reason for their existence, (p.2)

However, in contrast, he quotes the following words of Hyman (1973):

no single dogma is more central to the accepted philosophy of higher education than the notion that a university faculty member must be a scholar as well as a teacher, (p. 17)

Lahti (1973) has expressed concern over the neglect of the administrative role of the academic. He states that:

the primary source for filling key managerial positions (in higher education) is untrained, upwardly mobile, academicians who take their turn in the classroom and then become part of the higher education establishment, (p.34)

In his discussion of the administrative aspects of the academic's work, Miller (1974) has also stressed the importance of managerial capability. He states that:

The quality of management is the key to success in any institutionalised enterprise. (p.48)

In summary, one might tentatively conclude that academics must be capable of incorporating into their job the duties of teacher, researcher, consultant, and manager. Accordingly, it is possible that the reluctance on the part of academic institutions to reveal the priority areas for their staff has led to a neglect of important areas and/or an attempt to give equal attention to every facet of the job resulting in an unreasonable workload.

There is evidence to support both those suggestions. First, despite the increasing emphasis on research, a facet of academic life about which there is some measure of agreement amongst writers on the subject, Centra (1982) refers to a 1977 survey of U.S. professorate conducted by Ladd and Lipset which found that 29% had never published an article, 60% had never published more than four articles, and 59% had never written, or edited, a book or monograph. Secondly, Blackburn et al (1980) reported that studies consistently record academics working an average fifty hour week and, in the case of the University of California, a sixty two hour week.

The Specific Context of the Priorities Argument

The disagreement as to academic priorities was largely irrelevant to Hong Kong Polytechnic staff in the 1970s during which time the institution was teaching centred, focusing on professional and technical training. The 1980s heralded a move away from technical education and an increased emphasis on academic development, research, and consultancy. Whilst specific priorities were not set for members of staff, there was a general perception that the criteria for effective academic performance had changed; good classroom performance and a practical background in the area taught were no longer considered good enough. This perception was reinforced by appointments, at all levels, of candidates who possessed higher degree and research experience, and a promotion policy which placed increased emphasis on research output as the main criteria for advancement.

Hong Kong Polytechnic staff found themselves caught up in the priorities dilemma. To add to this, a number of new academic departments were introduced in the late 1980's, which resulted in the formation of the Department of Management from the¹ staff to the existing Department of Business and Management Studies. Hence, not only had Management staff to wrestle with the problem of priorities, but they also had to engage in forging an identity for a new Department. The BARS developmental procedure, in addition to acclimatising staff to performance appraisal, was seen as providing assistance in both the priorities and identity areas by encouraging staff to think about, and agree upon, priority work areas, or performance dimensions as they are termed in the BARS context.

BARS

Properties

BARS were first introduced by Smith and Kendall (1963) and promised to overcome some of the psychometric inadequacies associated with other rating formats. This promise was based on a rigorous procedure which produced a scale or set of scales containing critical incidents exemplifying levels of performance in the various aspects of a particular job. The developmental procedure sought to be objective and to incorporate a high measure of agreement on scale anchors amongst participants in the procedure. Participants were selected on the basis of their expertise gained from a close association with the job in question.

Since 1963, many studies have attempted to evaluate the psychometric properties of BARS (see for example Bernardin and Smith, 1981; Harrell and Wright, 1990; Kinicki et al. 1985; Wiersma and Latham, 1986). Interested readers are also referred to articles by Jacobs et al (1980), Kingstrom and Bass (1981), and Schwab et al (1975) which, when taken together, constitute a comprehensive review of individual and comparative studies examining the properties of BARS. These studies indicate that whilst BARS psychometric performance is not inferior to that of other formats, neither are BARS superior to other rating approaches.

Such results, however, need to be viewed in the light of methodological shortcomings associated with the comparative studies (Jacobs et al. 1980; Kingstrom and Bass, 1981; Schwab et al. 1975). First, almost all the comparative research has involved the evaluation of BARS relative to scales which have been developed using performance dimensions generated through BARS developmental procedures. Consequently, one has to question the extent to which one would expect to find meaningful differences in the psychometric properties of scales emanating from the same developmental procedure.

Secondly, comparative studies often operationalise psychometric properties differently. Hence when measures of leniency and inter-rater agreement for example, differ from study to study, it is difficult to reach meaningful conclusions as to the superiority, or otherwise, of one rating format over another.

Thirdly, in words of Kingstrom and Bass (1981):

comparison of rating formats are confounded

by differences in samples, organisations, types of raters, number of responses raters make per dimension, etc., from one study to another. (p.282)

Hence, comparative studies have been guilty of failing to compare like with like.

A final, particularly telling point, has been made regarding all the studies which attempt to compare the psychometric performance of BARS with other scales, namely, that in the absence of true scores for ratees, decisions on better or worse psychometric performance are very difficult. Hence, an apparent poorer performance of one format in comparison with another on, for example, the halo psychometric criterion may fail to take account of the fact that, in the case in question, dimensions of performance are, in actuality, highly interrelated.

In contrast with studies focusing on psychometric properties, evaluation of BARS against important qualitative criteria has produced more concrete findings in their favour. In order to illustrate this point, BARS performance is assessed here on the following bases: relevancy of the instrument, quality of feedback, and overall organisational benefits.

In general, any performance rating instrument should focus attention on essential job elements whilst excluding the non essential. In other words, items in the instrument should be relevant to the job in question. BARS are particularly strong in this area as performance dimensions and behavioural anchors are developed and agreed by experts in a particular job. Hence Jacobs et al (1980) have made the following assertion:

Perhaps the strongest attribute of the BARS methodology is its ability to yield job analysis information performed by the people who know the job best and written in that language, (p.606)

Ideally, evaluations should result in concrete, specific feedback. Specific feedback to the ratee, when it is positive, is likely to enhance job satisfaction and motivation and, as a result, future performance (Beatty, Schneier and Beatty, 1977). Conversely, specific feedback pointing out weaknesses can serve to indicate where improvement needs to be made and can lead to the development of training programmes (Blood, 1974). The BARS format, comprising specific behavioural examples, is likely to be effective in providing the type of feedback which is desirable. Some support for this view is found in a longitudinal study conducted by Ivancevich (1980) which showed that, in

comparison to a traits format, use of a behavioural expectation format (BARS) was associated with more favourable attitudes to performance appraisal, less job tension, more job satisfaction, greater organisational commitment and higher internal motivation.

The overall organisational benefits deriving from BARS are as a consequence of the instrument's specificity. It has been argued that, as BARS specifically indicate appropriate and inappropriate behaviours, it is possible that the feedback sessions of BARS appraisals could generate widespread discussion amongst employees leading to a climate in which informal goals are set aimed at improving organisation wide performance (Jacobs et al. 1980).

In addition, studies which have found significant disagreement in the rating of BARS items amongst organisational levels indicate the potential of BARS to stimulate discussion aimed at resolving differences amongst these levels over policies and priorities (Zedeck et al., 1974; Zedeck and Kafry, 1977).

Despite the benefits referred to above, BARS have been criticised for the time and cost involved in their development, largely due to the degree of participation required in the developmental process. This criticism has to be weighed against the fact that participation in scale construction can serve to focus staffs attention on important performance issues. Hence Campbell et al (1973) have made the following assertion:

It is our contention that most people in organisations seldom, if ever, give careful attention to what they really mean by effective performance. The (BARS development) procedure forces a confrontation with this question. Defining effective performance. . . . is an integral part of management, (p.32)

In summary, the psychometric superiority of BARS over other rating formats has not been clearly established and the scales can be time consuming and costly to develop. Nevertheless, BARS offer potential benefits which certainly merit the attention of organisations valuing rating scales which, at the developmental stage, force staff to consider the nature and constituent elements of effective performance, reflect the true nature of the jobs under evaluation, provide specific feedback, and facilitate organisation wide dialogue on goals, policies and priorities.

It was felt that the qualitative benefits described above were of particular value in a situation, such as

that prevailing in the Department of Management Studies of the Hong Kong Polytechnic, where the idea of implementing systematic performance appraisal for academic staff might raise apprehensions and objections. In such a situation, it was perceived that qualitative issues such as the relevancy and acceptability of the format to staff in general would be far more crucial to the success of any pilot study employing BARS, than would be the ability of BARS to perform well on less tangible and less immediate psychometric indices. Accordingly, in the Department of Management Studies, BARS were selected for development.

Developmental Procedure

The basic procedure for developing BARS is as follows:

- i) A group of experts, normally individuals who either supervise, or themselves carry out a job of work, list and define the various performance dimensions which constitute the job in question.
- ii) A second group of experts provides examples of good, average, and poor job behaviours and these examples are written in an expectation format.
- iii) A third group of experts is given a randomised list of job behaviours and is required to allocate these behaviours to performance dimensions. Job behaviours and performance dimensions are eliminated at this stage if there is no clear measure of agreement as to the behaviours which relate to the performance dimensions.
- iv) A fourth group of experts assigns numerical values to items according to a predetermined scale. The mean and variance of the values given to the behavioural examples are computed and items with small variances provide the anchors for the scale according to their mean values.

Ratees, Raters, and the Rating Method

The following questions are relevant to any proposed performance rating system:

- i) who should be rated?
- ii) who should do the rating?
and
- iii) how should the rating be carried out?

Head of Department as Ratee

A key question in the design of the appraisal system for the Department of Management was the extent to which it was desirable or feasible for the Department Head to be appraised by his staff. Cognisance had to be taken of the very different job nature of a Polytechnic head of department to that of academic members of staff responsible to him. A Polytechnic head of department's primary function was, at the time the project was undertaken, staff and resource management with little involvement in the typical day to day tasks of teaching staff. As the application of a BARS instrument of appraisal requires that ratees constitute a reasonably homogeneous group in terms of job content, it was decided to exclude the Head from the proposed appraisal system.

Students as Raters

Despite the disagreement as to academic priorities already discussed, there is little doubt amongst academics and writers on academic performance that teaching must figure as a critical aspect of an academic's job. It seemed logical that those persons best placed to evaluate this aspect of performance would be the recipients, the students. Research supports this view. (Remmers and Weisbrodt, 1965; McKeachie, 1969, Hildebrand and Wilson, 1970; Miller, 1974; Centra, 1982)

Colleagues as Raters

The inevitable question concerned the job areas or performance dimensions which colleagues were qualified to appraise. Factors to be accounted for were the dangers of 'mutual backscratching' or professional jealousy, particularly when colleague rating is to be used for merit or promotion decisions.

On balance, it appeared that colleague evaluation of teaching was of limited value as research suggests that it is likely to be based on pure supposition, (Centra, 1982) or merely reflects prior knowledge of student ratings (Murray, 1972).

Colleagues' assessment of performance dimensions other than teaching seemed to offer more promise. It was felt that such dimensions as research and administration were likely to be more appropriate vehicles for peer evaluation. Also, it was noted that, in the Polytechnic, individual members of staff were

apt to interact with different colleagues for different aspects of their performance, which implied that more than one member of staff would need to be involved in appraising the overall performance of a colleague.

Consequently, Centra's (1982) outline of a peer appraisal procedure was perceived to have merit particularly as first, it did not expose colleagues to 'backlash' over unpopular ratings and therefore limited the possibility of conflict arising within an academic department, and secondly, it allowed the ratee some discretion as to choice of raters.

The anonymous group procedure, proposed by Centra, allows the ratee to select five colleagues as potential raters. The departmental head selects three out of the five and then adds three more of his choice. Groups members are unaware of each other's identity and do not know who nominated them. The group never meets, so that no one evaluator can influence the rest of the group. Ratings are 'pooled' by the head to form an overall judgement of the ratee's performance.

Self Appraisal

Research from the field of management confirms the value of self appraisal. Mabe and West (1982) have shown that self evaluation may be a fair indication of actual, or future performance. They have also hinted at a link between the accuracy of self appraisal and level of intelligence, which indicated that self appraisal might be well suited to an academic institution such as the Polytechnic.

Shrauger and Osberg (1981) have asserted that self appraisal can be valuable because individuals possess a much larger database about themselves than any external assessor. In addition, Teel (1978) has disputed a commonly held notion, that self appraisal leads to self enhancement, and has proposed that underrating is more common.

The Multi-Appraisal Method

Stimson and Stokes (1980) have proposed an appraisal method which combines peer appraisal, self assessment, and the traditional superior-subordinate appraisal interview. The method requires the ratee to select five to eight raters including peers and subordinates. Raters are guaranteed anonymity and their ratings are summarised. Both superior and subordinate receive a copy of the summary rating. The subordinate then

completes a self assessment which is sent to the superior. A final appraisal form, reflecting the peer rating summary, ratee's self assessment, and superior's assessment of subordinate, provides the basis for the appraisal interview.

That method seemed to incorporate many of the preferred elements of performance appraisal. Anonymity of raters ensures that peers give a fair assessment although, given that the ratee has chosen potential raters, it was noted that total anonymity might not be assured. An acceptable compromise seemed to be that proposed by Centra (1982), and outlined earlier, in which an 'anonymous group' is constituted from a selection of the superior's nominees and those of the subordinate.

Procedure for the Development of a BARS Instrument in the Department of Management

The procedure employed for the development of BARS in the Department of Management is described below:

- i) All academic staff in the Department were requested to identify relevant performance dimensions.
- ii) Academic staff in the Department, and student groups, were asked to provide examples of good, average, and poor job behaviours for the dimensions identified in step (i). Student groups were asked only to provide examples for performance dimensions the results of which were within their experience e.g. teaching and counselling.
- iii) A list of performance dimensions and behavioural examples was forwarded to each member of academic staff of the Department with the request that behavioural examples be allocated to relevant performance dimensions. Where there was disagreement on the performance dimensions to which a particular behavioural example belonged (i.e. there was less than 70% agreement amongst participants in the exercise), the particular example was eliminated.
- iv) A list of the performance dimensions, together with the behavioural examples allocated to them as a result of step (ii) was then produced, copies sent to all academic staff and, in the case of student centred activities, to a student group, with the request that a rating be allotted to each of the behaviours, on a scale of 1 to 5, according to the individual's perception

as to the level of performance represented by the behavioural example. A numerical rating of 1 was to represent poor performance and a rating of 5 was to represent excellent performance.

- v) The mean and variance of the rating given to each behavioural example were computed and those items with smaller variances assigned as anchors for the rating scale in accordance with their mean values. As a result of the developmental process described above, it was possible to produce a BARS instrument covering the following performance dimensions: teaching, research, consultancy, administration (management), liaison, participation, and consulting. For reasons of space, only the scale constructed for teaching has been included in Appendix A to this paper. (The remaining scales are available from the author).

The above procedure departed from the original Smith and Kendall (1963) formula in using the same group of experts for each of the developmental stages. One reason for this departure was the small number of academic staff in the Department i.e. 30 in all, which did not allow for the formation of a number of different groups. Additionally, an important component of the BARS developmental procedure was the participation required of staff. In order to facilitate staff acceptance of a new performance appraisal system, the author chose to emphasise the participative element of the procedure the expense of strict adherence to the Smith and Kendall formula.

Proposed Rating Method for Academic Staff

The following rating method for staff was proposed, and in the context of developments in Hong Kong higher education referred to at the beginning of this paper, is suggested for present use:

- i) Members of staff select one class for which they have substantial teaching and counselling responsibility i.e. two performance dimensions identified in the developmental process (see previous section). The Head of Department, in consultation with relevant Course Leaders, arranges that one class contact session be set aside for the appraisal exercise. According to the date set by the Head, staff members issue appraisal forms, in the BARS format, to the classes in question. In order to preserve anonymity, no name is required on the form.

The forms are returned, on a pre-arranged date and time, to the Departmental Head via his Administrative Assistant.

- ii) Staff members each nominate five colleagues, either peers or subordinates, or both, as raters. The names are forwarded to the Head of Department who selects three from these nominees and adds three more of his choice. Nominees do not have to be confined to the Department of Management but can be drawn from any department, as the criterion for selection is the ability of the nominee to make an informed assessment, and such ability is not necessarily confined to a staff member's home department.

Accordingly, raters external to the Department assess ratees on the basis of their interaction with them in such activities as course committees, hosted by the external department, but in which ratees are active participants. It should be noted that all raters, including those external to the Department, are only called upon to rate those dimensions of performance that they have had an opportunity to observe through interaction with ratees.

- iii) The Head of Department arranges for the issue of appraisal forms (see Appendix A) to each of the six raters who are requested not to discuss their involvement in the appraisal exercise in order to ensure anonymity and influence-free assessment. The appraisal forms, which, as previously stated, allow raters to rate only those dimensions which they feel competent to judge, are returned, on a pre-determined date to the Head, via his Administrative Assistant. Similarly, individual ratees complete a self appraisal (Appendix A), and forward this to the Head. Each ratee keeps a copy of the self appraisal.

- iv) The Head arranges a date for an appraisal interview with individual ratees and ensures that both he, and the member of staff concerned, receive copies of the student and peer ratings, which include such behavioural examples as have been provided by the raters, before the appraisal interview.

- v) The appraisal interview takes place and focuses on differences between self appraisal and other appraisals. Both parties strive for agreement on a fair assessment of the ratee's performance, attempt to identify areas where improvement might be made, and decide on

appropriate courses of action to effect such improvement. A development plan, where agreed, and the date of any subsequent progress meeting, are noted on an appraisal interview/follow up form (see Appendix B). The form is signed by both the Head and the ratee and both are given a copy, a procedure which ensures that both parties are fully aware of the results of the appraisal interview.

Subsequent progress meetings will focus on the extent to which the development plan has been implemented thus far, its effects on ratees' performance, and the action required of the Head, or ratees themselves, to further the plan.

Due to the time and effort likely to be involved in the rating exercise, it was proposed that appraisals take place annually, although it was recommended that informal progress meetings between the Head of Department and each ratee be held at selected intervals. It was envisaged that such meetings would take place at intervals no greater than three months.

Updating the Rating Instrument

It was envisaged that the Administrative Assistant would extract from appraisal forms received, any new behavioural examples provided by raters and store them in computer memory categorised by performance dimension. On completion of one round of staff appraisal it was proposed that the new behavioural examples be converted into the expectation format and listed in random order together with the performance dimensions. The list would then be forwarded to members of academic staff, and to selected student groups where applicable, and steps (iii), (iv), and (v) of the procedure for the development of the BARS instrument (described above) repeated, for the purpose of expanding, and updating the anchors.

Implementation

It was recognised that the introduction of a performance appraisal system into an academic department was likely to generate a degree of apprehension amongst staff. Furthermore, it was acknowledged that successful implementation was likely to depend on the extent to which apprehension was prevented from being translated into change resistant behaviour.

In view of this, a review of some of the

literature on the management of change was carried out.

The Management of Change

Kirkpatrick (1985) has identified three components of effective change management: empathy, communication, and participation.

Empathy is seen as an important attribute of the manager involved in change. It consists of the ability to view change from the point of view of those affected and requires that managers make efforts to get to know their staff and anticipate reactions to change.

Kirkpatrick has stressed the importance of two-way communication between manager and staff, as far in advance of the change as possible, to give people time to adapt to new ideas and proposals. Additionally, he believes that continuous dialogue should take place throughout implementation and until the change has been effected. It is equally important to maintain this dialogue after the change has taken place in order to monitor its effect on staff and their reaction to the new status quo, and to gather information for possible fine tuning.

Participation means getting people involved so that they take ownership of the change. Participation facilitates acceptance and enhances the quality of change through improved input of ideas.

In a similar vein, Kotter and Schlesinger (1979) have identified three key strategies for introducing change. These are:

- a) education, and communication
- b) participation and involvement, and;
- e) facilitation and support.

Education, and communication involve explaining as much as possible about the nature of the change and its possible effects, to those affected. As with Kirkpatrick (1985), the importance of two-way communication is stressed, this time with an emphasis on generating understanding.

Participation and involvement help unfreeze fixed attitudes, foster trust between manager and subordinates, and help the organisation to benefit from the increased contribution of ideas.

Facilitation and support is associated with empathy. It involves not only listening to fears and anxieties but also providing time off and/or training where necessary.

Stewart (1983) has proposed a number of strategies for introducing change. She has recommended that the reasons for change and the

anticipated process of change should be communicated to those involved and that change be initiated first as a pilot study so that initial uncertainties and resistance can be overcome, and potential problem areas identified. She further suggests that peer group pressure be used to facilitate the process of change. This involves piloting a scheme with people whom other members of the organisation are likely to use as a reference group. She has also stressed the importance of the incremental approach which allows time for people to adapt to change.

In summary, core change strategies involve empathy, participation, and communication. The incremental approach should be considered and managers should make use of pilot schemes and peer group pressure to smooth the implementation process.

The Way Forward Proposed for the Department of Management of the Hong Kong Polytechnic

In order to place the following recommendations in context, it is useful to note that in order to overcome inevitable staff reluctance to embrace the idea of appraisal, it was necessary to emphasise continually the exploratory nature of this initiative, both formally and informally.

It was communicated to staff that the process of agreeing on critical work areas and developing examples of performance was of value in itself, particularly for the purpose of forging an identity for a newly formed department. It was also emphasised that, should the resulting appraisal instrument be considered a suitable starting point for moving towards a fully fledged performance appraisal system based on BARS, such a move would only take place with staff agreement.

It was against this background that the following recommendations for furthering the initiative in the Department of Management of the Hong Kong Polytechnic were made.

i) The author would produce a report on this initiative which would be issued to all members of academic staff within the Department. The report would be accompanied by a recommendation that it form the basis for both formal and informal discussions, aimed at establishing whether or not the notion of staff appraisal was acceptable to members of staff. In addition, assuming that most staff would not object to some form of appraisal, an action plan for implementation on a trial basis should

be agreed. The action plan should address, in particular, the issue of preparatory staff training.

At this point, the Head should indicate a proposed time frame for introduction which would allow a twelve month period for discussion before initial implementation.

In the context of management of change literature, the recommended action was viewed as initiating open communications on the proposal and allowing for participation in the implementation plan. Equally, the undertaking not to implement a system for twelve months was seen to be akin to an incremental approach which allows time for the necessary attitudinal adjustments to take place.

It was stressed that the Head should emphasise the trial run nature of the proposed implementation to prevent the hardening of attitudes associated with change perceived by those affected as irreversible, or incapable of modification.

ii) The Head should issue an open invitation to all members of academic staff in the Department to discuss with him, on a personal basis, any problems they might have, or suggestions for improvement they may wish to make regarding the proposed system.

The importance of empathy has already been noted, and it is worth emphasizing that at these meetings, the Head should be prepared to listen, allay fears and anxieties where possible, provide any support he may feel is appropriate, and make adjustments to the proposed system where necessary.

iii) The Head should make efforts to enlist voluntary support for a pilot scheme. Preferably the pilot scheme should involve one or more of the Department's Course Leaders, (i.e. staff members responsible for the management and administration of academic programmes). The knowledge gained from the pilot scheme should be used to inform Department discussions of the problems (if any) encountered in the system, but not previously anticipated, and of any modifications which might be desirable prior to implementation.

It was felt that the pilot scheme would help the Head to decide on the practicality and feasibility of the system and the extent to which any actual, or potential opposition might be reduced, or removed as a result of

modifications. Additionally, it was felt that the results of the pilot scheme would form a useful starting point for any training deemed necessary for the successful introduction of the scheme.

The involvement of Course Leaders in the pilot scheme was seen to be particularly useful, as this group had recently been afforded the status of key individuals in the institution by Polytechnic management and had become a visible reference group for other members of staff.

Discussion

It was recognised that the above mentioned implementation strategy did not guarantee staff acceptance of the performance appraisal system. Additionally, during the course of the initiative, the author was fully aware of the possibility of opposition, amongst staff, to the notion of appraising academic performance. However the response of staff in the BARS developmental procedure was encouraging and something upon which to build.

It was felt that at the implementation stage, it would be necessary to stress constantly the value of the proposed system for members of staff in enabling them to effectively design their own appraisal instrument which would emphasise aspects of performance perceived by them to be important. This would seem to be infinitely preferable to having an instrument imposed on them from above.

In the event that the proposed appraisal system was rejected outright, the author was convinced that the exercise would still have been useful as a means of facilitating clarification of, and agreement on, work priorities within the Department and stimulating discussion amongst staff on the subject of academic performance appraisal.

Regarding the instrument, as anticipated, the Smith and Kendall (1963) procedure as modified in this initiative required considerable time and effort on the part of otherwise busy staff to develop BARS. However, informal discussions with members of staff suggested that the issue of time and cost should not be overemphasised as the developmental procedure was seen to be beneficial in focusing attention on job performance and generating an interest in performance appraisal.

Given current developments in Hong Kong referred to earlier, it seems that the time is right

either for the system described in this paper to be implemented, or for similar initiatives aimed at fair and objective appraisal to be carried out in UPGC funded institutions.

References

- Beatty, R.W., Schneier, C.E., & Beatty, J.R. (1977). An Empirical Investigation of Perceptions of Ratee Behaviour Frequency and Ratee Behaviour Change Using behavioural Expectation Scales (BES). *Personnel Psychology*, 30: 647-658.
- Bernardin, H.J., & Smith, P.C. (1981). A Clarification of Some Issues Regarding the Development and Use of Behaviourally Anchored Rating Scales (BARS). *Journal of Applied Psychology*, 66: 458-463.
- Blackburn, R.T., O'Connell, C., & Pellino, G. (1980). Evaluating Faculty Performance. In A. Jedamus, M.V. Peterson, & Associates (Eds.). *Improving Academic Management*. San Francisco, Jossey Bass.
- Blood, M.R. (1974). Spin-offs from Behavioural Expectation Scale Procedures. *Journal of Applied Psychology*, 59: 513-515.
- Campbell, J.P., Dunnette, M.D., Arvey, R.D., & Hellervik, L.V. (1973). The Development and Evaluation of Behaviourally Based Rating Scales. *Journal of Applied Psychology*, 57: 15-22.
- Caplow, T., & McGee, R.T. (1961). *The Academic Market Place*. New York, New York Science.
- Centra, J.A., (1982). *Determining Faculty Effectiveness*. San Francisco, Jossey Bass.
- Greenaway, H., & Harding A.G. (1978). *The Growth of Policies for Staff Development*. University of Surrey Society for Research in Higher Education.
- Harrell, A., & Wright, A. (1990). Empirical Evidence on the Validity and Reliability of Behaviourally Anchored Rating Scales for Auditors. *Auditing* 9(3): 134-139.
- Hildebrand, M., & Wilson, R.C. (1970). *Effective University Teaching and Its Evaluation*. Berkeley, Calif, Centre for Research in Higher Education, 1970.
- Hyman, S.C. (1973). The Teacher as Scholar Ersatz Image? *The Chronicle of Higher Education*, January 8.
- Ivancevich, J.M. (1980). Behavioural Expectation Scales versus Non-Anchored and Trait Rating Systems: A Sales Personnel Application. *Applied Psychological Measurement*, 4: 131-134.
- Jacobs, R., Kafry, D., & Zedeck, S. (1980). Expectations of Behaviourally Anchored Rating Scales. *Personnel Psychology*, 33: 595-639.
- Kingstrom, P.O., & Bass, A.R. (1981). A Critical Analysis of

Towards a Behaviourally Based Performance Appraisal System

- Studies Comparing Behaviourally Anchored Rating Scales (BARS) and other Rating Formats. *Personnel Psychology*, 34: 263-289.
- Kinicki, A.J., Bannister, B.D., Horn, P., & De Nisi, A.S. (1985). Behaviourally Anchored Rating Scales vs Summated Rating Scales: Psychometric Properties and Susceptibility to Rating Bias. *Educational, and Psychological Measurement*, 45: 535-549.
- Kirkpatrick, D.L. (1985). *How to Manage Change Effectively*. San Francisco, Jossey Bass.
- Kotter, J.P., & Schlesinger, L.A. (1979). Choosing Strategies for Change. *Harvard Business Review*, March-April, 106-114.
- Ladd, E.C., & Lipset, S.M. (1977). Scholarly Articles Published by Professors. *The Chronicle of Higher Education*, November 28.
- Lahti, R.E. (1973). *Innovative College Management*. San Francisco, Jossey Bass.
- Landy, F.J., & Fair, J.L. (1983). *The Measurement of Work Performance*. New York, Academic Press.
- Mabe, P.A., & West, S.G. (1982). Validity of Self Evaluation of Ability: A Review and Meta-analysis, *Journal of Applied Psychology*, 67: 280-296.
- McKeachie, W.J. (1969). Student Ratings of Faculty. *AAUP Bulletin*, 55: 439-444.
- Miller, R.I. (1974). *Evaluating Faculty Performance*. San Francisco, Jossey Bass.
- Miller, R.I. (1975). *Developing Programmes for Faculty Evaluation*. San Francisco, Jossey Bass.
- Murray, H.G. (1972). The Validity of Student Rating of Teaching Ability. Paper Presented at the Canadian Psychological Association meeting, Montreal.
- Remmers, H.H., & Weisbrodt, J.A. (1965). *Manual of Instruction for the Purdue Rating Scale for Instruction*. West Lafayette, Indiana, University Book Store.
- Schwab, D.P., Heneman, H.G., & De Cotiis, T.A. (1975). Behaviourally Anchored Rating Scales: A Review of the Literature. *Personnel Psychology*, 28: 549-562.
- Shrauger, J.S., & Osberg, T.M. (1981). The Relative Accuracy of Self Predictions and Judgements by Others in Psychological Assessment. *Psychological Bulletin*, 90: 322-351.
- Smith, P.C., & Kendall, L.M. (1963). Retranslation of Expectations: An Approach to the Construction of Unambiguous Anchors for Rating Scales. *Journal of Applied Psychology*, 47: 149-155.
- Stewart, V. (1983). *Change: The Challenge For Management*. Maidenhead, McGraw-Hill.
- Stimson, J., & Stokes, J. (1980). How to Multi-Appraise. *Management Today*, June, 43-53.
- Teel, K.S. (1978). Self Appraisal Revisited. *Personnel Journal*, 57: 364-367.
- Wiersma, U., & Latham, G. (1986). The Practicality of Behavioural Observation Scales, Behavioural Expectation Scales and Trait Scales. *Personnel Psychology*, 39: 619-628.
- Zedeck, S., Imparato, N., Krausz, M., & Oleno, T. (1974). Development of Behaviourally Anchored Rating Scales as a Function of Organisational Level. *Journal of Applied Psychology*, 59: 249-252.
- Zedeck, S., & Kafry, D. (1977). Capturing Rater Policies for Processing Evaluation Data. *Organisation Behaviour and Human Performance*, 18: 269-294.

POUNDER J.S.

APPENDIX A

HONG KONG POLYTECHNIC DEPARTMENT OF MANAGEMENT

ACADEMIC STAFF PERFORMANCE APPRAISAL/SELF APPRAISAL FORM

RATEE'S NAME- _____

Instructions to Rater (Other than Self-Rater)

This form contains behavioural examples on a numerical scale, for the various dimensions of performance which constitute the job of a member of academic staff in the Department of Management.

The behavioural examples are numerically scaled to indicate the type of job behaviour which might be expected at different levels of competence in the various performance dimensions.

Your task is to give a numerical rating to the ratee's job performance, with reference to the behavioural examples provided. In other words, you are required to identify those behavioural examples which, in your opinion, most closely match the typical behaviour you would expect from the ratee in the performance dimensions concerned, and allocate a numerical rating with reference to the numerical scale provided.

Note that scale ranges from 1 to 5, and is graduated in steps of 0.5. Therefore the numerical rating which you should enter in the space provided, should be either 1, 1.5, 2, 2.5, etc up to 5.

Should you feel that the behavioural examples provided in the form are incomplete or unrepresentative of the ratee's performance and that you are in position to offer a more appropriate example, please enter this example in the space provided and allocate the numerical rating which you feel represents the level of performance of the ratee as illustrated by your example. Comparing your example with the examples already provided in the form should help you to decide on an appropriate numerical rating.

You should note that each of the sections of the appraisal form is devoted to one dimension of performance and that there is a space provided for rater's own example, and numerical rating, in each section.

You need only offer a rating for those dimensions of the ratee's performance which you feel competent to judge, through your observation of, and/or interaction with the ratee (self raters should rate all performance dimensions).

If you are a member of the student body you are only required to rate the performance dimensions: teaching and counselling.

This form should be returned to the Administrative Assistant (Management) no later than _____.

Instructions to Self Rater

Please refer to the above section and follow the instructions where appropriate. You should provide a rating for all dimensions of your performance with reference to the behavioural example in each dimension which you feel is most representative of your typical performance.

Should you feel that all the behavioural examples in one or more of the performance dimension scales are unrepresentative of your performance, please provide your own example in the designated space and allocate the numerical rating which you feel is represented by this example. Comparing your example with the other examples provided in the form should help you to decide on an appropriate numerical rating.

This form should be returned to the Administrative Assistant (Management) no later than _____.


Section 1

Performance Dimension: Teaching

The preparation and delivery of current student centred materials; the design of a varied learning session including the structuring of student learning experiences; imparting up to date knowledge, experience and opinion in such a way as to facilitate the learning process; helping students to make sense of each learning session; providing fair and informative feedback to students on their performance.

Performance which exceeds standard	r	
Performance around standard		<p>4.5 <input type="checkbox"/> Could be expected to utilise seminars and tutorials in such a way that they become effective learning sessions. <input type="checkbox"/> Could be expected to adopt a challenging and interactive style of presentation.</p>
		<p>4 <input type="checkbox"/> Could be expected to encourage questions and seek for multi-solution answers. <input type="checkbox"/> Could be expected to deliver knowledge and skills which are likely to be relevant to a student's future career. <input type="checkbox"/> Could be expected to provide detailed, actionable feedback on student performance. <input type="checkbox"/> Could be expected to illustrate the application of theory with reference to practical examples drawn either from his/her own work experience, or from the experiences of students.</p>
		<p>3.5 <input type="checkbox"/> Could be expected to have explained and clarified the objectives of the subject being taught, to the students. <input type="checkbox"/> Could be expected to explain the relationship of the subject to the course in general. <input type="checkbox"/> Could be expected to produce high quality examination papers for the subjects in which he/she is involved. <input type="checkbox"/> Could be expected to make clear his/her expectations of students in relation to the course he/she is teaching.</p>
		<p>3 <input type="checkbox"/> Could be expected to use materials which are appropriate and relevant to the subject. <input type="checkbox"/> Could be expected to prepare, and issue course and subject related handouts. <input type="checkbox"/> Could be expected to recommend references appropriate to the subject.</p>
		<p>2.5 <input type="checkbox"/> Could be expected to have given a general introductory lecture related to the subject being taught, but not specifically explained and clarified objectives. <input type="checkbox"/> Could be expected to produce examination papers for his/her subject which, whilst serving the purposes of assessment, display little imagination or inventiveness.</p>

Performance Dimension: Teaching

Performance in need of improvement to meet standard		2	_____	Could be expected to make little attempt to vary the method of presentation.
		1	_____	Could be expected to provide little opportunity for student questions relating to the topic presented.
		1	_____	Could be expected to use materials which have not been updated for some time.
		1.5	_____	Could be expected to make little effort to explain or clarify objectives or to generally introduce the subject.
			_____	Could be expected to exert little effort to make seminars and tutorials, effective and relevant learning experiences for students.
			_____	Could be expected to pay scant attention to the quality, level, and relevance of the examination papers for which he/she has prime responsibility.
		1	_____	Could be expected to deliver a lecture which consists of reading his/her notes or reciting from a book.
			_____	Could be expected to present theories without reference to practical examples.

Rater's own behavioural example illustrating expected behaviour of ratee in this dimension of performance (optional____see rater's instructions);

____Numerical Rating

APPENDIX B

**HONG KONG POLYTECHNIC
DEPARTMENT OF MANAGEMENT**

**ACADEMIC STAFF PERFORMANCE APPRAISAL
INTERVIEW/FOLLOW UP FORM**

RATEE'S NAME: _____

Instructions for Completion

This form should be completed during the appraisal interview by the Head of Department. On its completion, this form must be signed by the Head of Department and the ratee and both parties given a copy.

Developmental Areas Identified

Development Plan

Date of Next Meeting

Head of Department's Signature

Ratee's Signature

J.S. POUNDER is Head of Department of Management, Lingnan College Hong Kong.